Figure 1(A)

G*GATC*GTCTCAGGTCAG<u>CGGAGGGA</u>   25
                     SL33
<u>GGAGACTTA</u>TAGACCTATCCAGTCT   50
TCAAGGTGCTCCAGAAAGCAGGAGT   75
TGAAGACCTGGGTGTGAGGGACACA   100
TACATCCTAAAAGCACCACAGCAGA   125
GGAGGCCCAGGCAGTGCCAGGAGTC   150
AAGGTTCCCAGAAGACAAACCCCCT   175
AGGAAGACAGGCGACCTGTGAGGCC   200
<u>CTAGAGCACCACCTTAA</u>GAGAAGAA   225
       SL34
GAGCTGTAAGCCGGCCTTTGTCAGA   250
GCCATC**ATG**GGGGACAAGGATATGC   275
CTACTGCTGGGATGCCGAGTCTTCT   300
CCAGAGTTCCTCTGAGAGTCCTCAG   325
AGTTGTCCTGAGGGGGAGGACTCCC   350
AGTCTCCTCTCCAGATTCCCCAGAG   375
TTCTCCTGAGAGCGACGACACCCTG   400
TATCCTCTCCAGAGTCCTCAGAGTC   425
GTTCTGAGGGGGAGGACTCCTCG**GA**   450
TCCTCTCCAGAGACCTCCTGAGGGG   475
AAGGACTCCCAGTCTCCTCTCCAGA   500
TTCCCCAGAGTTCTCCTGAGGGCGA   525
CGACACCCAGTCTCCTCTCCAGAAT   550
TCTCAGAGTTCTCCTGAGGGGAAGG   575
ACTCCCTGTCTCCTCTAGAGATTTC   600
TCAGAGCCCTCCTGAGGGTGAGGAT   625
GTCCAGTCTCCTCTGCAGAATCCTG   650
CGAGTTCCTTCTTCTCCTCTGCTTT   675
ATTGAGTATTTTCCAGAGTTCCCCT   700

Figure 1(B)

```
GAGAGAACTCAGAGTACTTTTGAGG    725
GTTTTCCCCAGTCTCCTCTCCAGAT    750
TCCTGTGAGCTCCTCCTCCTCCTCC    775
ACTTTATTGAGTCTTTTCCAGAGTT    800
CCCCTGAGAGAACTCAGAGTACTTT    825
TGAGGGTTTTCCCCAGTCTCTTCTC    850
CAGATTCCTATGACCTCCTCCTTCT    875
CCTCTACTTTATTGAGTATTTTCCA    900
GAGTTCTCCTGAGAGTGCTCAAAGT    925
ACTTTTGAGGGTTTTCCCCAGTCTC    950
CTCTCCAGATTCCTGGGAGCCCCTC    975
CTTCTCCTCCACTTTACTGAGTCTT    1000
TTCCAGAGTTCCCCTGAGAGAACTC    1025
ACAGTACTTTTGAGGGTTTTCCCCA    1050
GTCTCCTCTCCAGATTCCTATGACC    1075
TCCTCCTTCTCCTCTACTTTATTGA    1100
GTATTTTCCAGAGTTCTCCTGAGAG    1125
TGCTCAAAGTACTTTTGAGGGTTTT    1150
CCCCAGTCTCCTCTCCAGATTCCTG    1175
GGAGCCCCTCCTTCTCCTCCACTTT    1200
ACTGAGTCTTTTCCAGAGTTCCCCT    1225
GAGAGAACTCACAGTACTTTTGAGG    1250
GTTTTCCCCAGTCTCCTCTCCAGAT    1275
TCCTATGACCTCCTCCTTCTCCTCT    1300
ACTTTATTGAGTATTTTACAGAGTT    1325
CTCCTGAGAGTGCTCAAAGTGCTTT    1350
TGAGGGTTTTCCCCAGTCTCCTCTC    1375
CAGATTCCTGTGAGCTCCTCTTTCT    1400
```

Figure 1(C)

```
CCTACACTTTATTGAGTCTTTTCCA    1425
GAGTTCCCCTGAGAGAACTCAGAGT    1450
ACTTTTGAGGGTTTTCCCCAGTCTC    1475
CTCTCCAGATTCCTGTGAGCTCCTC    1500
CTCCTCCTCCTCCACTTTATTGAGT    1525
CTTTTCCAGAGTTCCCCTGAGTGTA    1550
CTCAAAGTACTTTTGAGGGTTTTCC    1575
CCAGTCTCCTCTCCAGATTCCTCAG    1600
AGTCCTCCTGAAGGGGAGAATACCC    1625
ATTCTCCTCTCCAGATTGTTCCAAG    1650
TCTTCCTGAGTGGGAGGACTCCCTG    1675
TCTCCTCACTACTTTCCTCAGAGCC    1700
CTCCTCAGGGGGAGGACTCCCTATC    1725
TCCTCACTACTTTCCTCAGAGCCCT    1750
CCTCAGGGGGAGGACTCCCTGTCTC    1775
CTCACTACTTTCCTCAGAGCCCTCA    1800
GGGGGAGGACTCCCTGTCTCCTCAC    1825
TACTTTCCTCAGAGCCCTCCTCAGG    1850
GGGAGGACTCCATGTCTCCTCTCTA    1875
CTTTCCTCAGAGTCCTCTTCAGGGG    1900
GAGGAATTCCAGTCTTCTCTCCAGA    1925
GCCCTGTGAGCATCTGCTCCTCCTC    1950
CACTCCATCCAGTCTTCCCCAGAGT    1975
TTCCCTGAGAGTTCTCAGAGTCCTC    2000
CTGAGGGGCCTGTCCAGTCTCCTCT    2025
CCATAGTCCTCAGAGCCCTCCTGAG    2050
GGGATGCACTCCCAATCTCCTCTCC    2075
AGAGTCCTGAGAGTGCTCCTGAGGG    2100
```

Figure 1(D)

GGAGGATTCCCTGTCTCCTCTCCAA   2125
ATTCCTCAGAGTCCTCTTGAGGGAG   2150
AGGACTCCCTGTCTTCTCTCCATTT   2175
TCCTCAGAGTCCTCCTGAGTGGGAG   2200
GACTCCCTCTCTCCTCTCCACTTTC   2225
CTCAGTTTCCTCCTCAGGGGGAGGA   2250
CTTCCAGTCTTCTCTCCAGAGTCCT   2275
GTGAGTATCTGCTCCTCCTCCACTT   2300
CTTTGAGTCTTCCCCAGAGTTTCCC   2325
TGAGAGTCCTCAGAGTCCTCCTGAG   2350
GGGCCTGCTCAGTCTCCTCTCCAGA   2375
GACCTGTCAGCTCCTTCTTCTCCTA   2400
CACTTTAGCGAGTCTTCTCCAAAGT   2425
TCCCATGAGAGTCCTCAGAGTCCTC   2450
CTGAGGGGCCTGCCCAGTCTCCTCT   2475
CCAGAGTCCTGTGAGCTCCTTCCCC   2500
TCCTCCACTTCATCGAGTCTTTCCC   2525
AGAGTTCTCCTGTGAGCTCCTTCCC   2550
CTCCTCCACTTCATCGAGTCTTTCC   2575
AAGAGTTCCCCTGAGAGTCCTCTCC   2600
AGAGTCCTGT**GAT**CTCCTTCTCCTC   2625
CTCCACTTCATTGAGCCCATTCAGT   2650
GAAGAGTCCAGCAGC<u>CCAGTAGATG</u>   2675
<div align="center">SL26</div>
<u>AATATACAAGTT</u>CCTCAGACACCTT   2700
GCTAGAGAGTGATTCCTTGACAGAC   2725
AGCGAGTCCTTGATAGAGAGCGAGC   2750
CCTTGTTCACTTATACACTGGATGA   2775
AAAGGTGGACGAGTTGGCGCGGTTT   2800

Figure 1(E)

```
CTTCTCCTCAAATATCAAGTGAAGC      2825
      SL27
AGCCTATCACAAAGGCAGAGATGCT      2850
GACGAATGTCATCAGCAGGTACACG      2875
GGCTACTTTCCTGTGATCTTCAGGA      2900
AAGCCCGTGAGTTCATAGAGATACT      2925
TTTTGGCATTTCCCTGAGAGAAGTG      2950
GACCCTGATGACTCCTATGTCTTTG      2975
TAAACACATTAGACCTCACCTCTGA      3000
GGGGTGTCTGAGTGATGAGCAGGGC      3025
ATGTCCCAGAACCGCCTCCTGATTC      3050
TTATTCTGAGTATCATCTTCATAAA      3075
GGGCACCTATGCCTCTGAGGAGGTC      3100
ATCTGGGATGTGCTGAGTGGAATAG      3125
GGGTGCGTGCTGGGAGGGAGCACTT      3150
TGCCTTTGGGGAGCCCAGGGAGCTC      3175
CTCACTAAAGTTTGGGTGCAGGAAC      3200
ATTACCTAGAGTACCGGGAGGTGCC      3225
CAACTCTTCTCCTCCTCGTTACGAA      3250
TTCCTGTGGGGTCCAAGAGCTCATT      3275
CAGAAGTCATTAAGAGGAAAGTAGT      3300
AGAGTTTTTGGCCATGCTAAAGAAT      3325
ACCGTCCCTATTACCTTTCCATCCT      3350
CTTACAAGGATGCTTTGAAAGATGT      3375
GGAAGAGAGAGCCCAGGCCATAATT      3400
GACACCACAGATGATTCGACTGCCA      3425
CAGAAAGTGCAAGCTCCAGTGTCAT      3450
GTCCCCCAGCTTCTCTTCTGAGTGA      3475
AGTCTAGGGCAGATTCTTCCCTCTG      3500
```

Figure 1(F)

AGTTTGAAGGGGGCAGTCGAGTTTC	3525
TACGTGGTGGAGGGCCTGGTTGAGG	3550
CTGGAGAGAACACAGTGCTATTTGC	3575
ATTTCTGTTCCATATGGGTAGTTAT	3600
GGGGTTTACCTGTTTTACTTTTGGG	3625
TATTTTTCAAATGCTTTTCCTATTA	3650
ATAACAGGTTTAAATAGCTTCAGAA	3675
TCCTAGTTTATGCACATGAGTCGCA	3700
CATGTATTGCTGTTTTTCTGGTTTA	3725
AGAGTAACAGTTTGATATTTTGTAA	3750
AAACAAAAACACACCCAAACACACC	3775
ACATTGGGAAAACCTTCTGCCTCAT	3800
TTTGTGATGTGTCACAGGTTAATGT	3825
GGTGTTACTGTAGGAATTTTCTTGA	3850
AACTGTGAAGGAACTCTGCAGTTAA	3875
ATAGTGGAATAAAGTAAAGGATTGT	3900
TAATGTTTGCATTTCCTCAGGTCCT	3925
TTAGTCTGTTGTTCTTGAAAACTAA	3950
AGATACATACCTGGTTTGCTTGGCT	3975
TACGTAAGAAAGTAGAAGAAAGTAA	4000
ACTGTAATAAATAAAAAAAAAAAAA	4025
AAAAAA	4031

FIG. 2(A)

```
A1  ⌐exon I                                                                                              exon I⌐ intron I
C1    CCATTCTGAGGGACGGGCGTA GAGTTCGGCGGCCGAAGGAACCT GACCCAGGCTCTGTGAGGAG GCAAG⌐g⌐...                         27
    ..................                                                          GGATCGT  CTCAGGTCAGCGGGAGGAGG ......

A1  acagagagaggaggatgcacag ggtgtgccagcagtgaatgt tt----...gccctgaatgca caccaaggggccccacctgcc acaggacacataggactca
C1  CCACAGAGGAGGAGGCCCAG  GCAGTGCCAGGAGAGTCAAG⌐g⌐t gagtgcacgacctgactgtg taccaagggccgtacccca  gaaacagtgtcagacctggc      115
                          exon II⌐intron II

A1  cagagtctggcctcacctcc ctactgtcgtcctgtagaa tcgac-ctctgctggccggc tgtaccctga-gtaccctct cacttcctccttcag⌐GTTTT
C1  agcaccggccctgtagccac ccactgtcattcctggtgcc tcatggctctgcctgccagc tgtgcccgagggtgcttct cgcgtccttctacag⌐GTTCC      158
                                                                                          intron II⌐exon II

A1  CAGGGGACAGGGCCAACCCAG AGGACAGGAGGATTCCCTGGAGG  CCACAGAGAGGAGGCACC----A AGGAGAAGAAGATCT⌐gtaagtag gccttgttagagtctccaa
C1  CAGAAGACAAACCCCCTAGG  AAGACAGGCGACCTGTGAGG  CCCTAGAGCACCACCTTAAG AGAAGAAGAGCTGTAAGCCG GCCTTTGTCAGAGCCATCAT      258
                                                                          exon II⌐intron II                      M

A1  ggttcag--ttctcagctgag gcctctcacacactccctct ctccc-cag⌐GCCTGTGGGGTC TTCATTG-CCCAGCTCCTGC CCACACTCCTGCCCTGCTGCC
C1  G⌐G⌐gtgagtttctcagctgag gccactggcactgtccctct ctcccctcagtcctggggat cccatcatacctattcgtgt tcacacgttacctgctgct      188
exon III⌐G⌐intron III                                      intron III⌐exon III

                         M   S   L   E   Q   R   S   L   H   C   K   P   E   E   A   L   E   A   Q   Q   E   A   L   G   L   V   C   V
A1  CT--GACGAGAGTCATCATG TCTCTTGAGCAGAGGAGTCT GCACTGCCAAGCCTGAGGAAG CCCTTGAGGCCCAACAAGAG GCCCTGGGGCTGGTGTGTGT   28
C1  cctgaacaatatcatcatg cctctcttctaaacccttcc acgcccagcttgagcaag gcttccagaaggcaattttc atactggagttggtagatgc     286

                         Q   A   A   T   S   S   S   P   L   V   L   G   T   L   E   E   V   P   T   A   G   S   T   D   P   P   Q   S   P   Q   G   A
A1  GCAGGCTGCCACCTCCTCCT CCTCTCCTCTGGTCCTGGGC ACCCTGGAGGAGGTGCCCAC TGCTGGGGTCAACAGGATCCTC CCCAGAGTCCTCAGGGGAGCC   61
C1  agaggatcccca--                                                                                              386

    S   A   F   P   T   T   I   N   F   T   R   Q   R   Q   P
A1  TCCGCCTTTCCCACTACCAT CAACTTCACTCGACAGAGGC AACCC-                                                             76
C1  ..................                                                                                          431
```

FIG. 2(B)

```
A1   tcccctcgtcctcctcctgt ttttcttctcatcctcatc ctcctctctgcttctgcgtt ctccagGGGACAAAGGATATG CCTACTGCTGGGATGCCGAG   294
C1                                                              intron III|          D  K  D  M  P  T  A  G  M  P  S      13
                                                                        exon IV

A1   TCTTCTCCAGAGAGTCCTCTG AGAGTCCTCAGAGAGTTGTCCT GAGGGGGGAGGAGGACTCCCAGTC TCCTCTCCAGAGATTCCCCCAGA GTTCTCCTGAGAGAGGCGACGAC   394
C1    L  L  Q  S  S  E    S  P  Q  S |Y  P  E  G  E  D  S  Q  S  P  L  Q  I  P| |Q  S  P  E  S  D|                          46

A1   ACCCTGTATCCTCTCCAGAG TCCTTCAGAGTCGTTCTGAGG GGGAGGAGGACTCCTCGGATCCT CTCCAGAGACCTCCTGAGGGG GAAGGAGGACTCCCAGTCTCCTC   494
C1    T  L  Y  P  L  Q  S |P  Q  S  R  S  E  G  E  D  S  S  D  P  L  Q  R| |P  P  E  G  K  D  S  Q  S  P  L                80

A1   TCCAGATTCCCCCAGAGTTCT CCTGAGGGGGGCGACGACACACCCA GTCTCCTCTCCAGAGAATTCTC AGAGTTCTCCTGAGAGGGGAAG GACTCCCTGTCTCCTCCTCTAGA   594
C1   |Q  I  P| Q  S  S  P  E  G  G  D  D  T  Q  S  P  L  Q  N  S  Q  S  S  P  E  G  K  D  S  L  S  P  L  E                   113

A1   GATTTCTCAGAGCGCCTCCTG AGGGTGAGGAGGATGTCCAGTCT CCTCTGCAGAATCCTGCGAG TTCCTTCTTCTTCTCCTGCTT TATTGAGTAGTATTTCCAGAGT   694
C1   |I  S| Q  S  P  P  E  G  E  D  V  Q  S  P  L  Q  N  P |A  S  S  F  F  S  S  A  L  L  S  I  F  Q  S|                     146

A1   TCCCCTGAGAGTATTCAAAG TCCTTTTGAGGGGTTTTCCCC AGTCTGTTCTCCAGATTCCT GTGAGCGCGCCGCCTCCTCCTC CACTTTAGTGAGTATTTTCC   794
C1   |S  P  E  S  I  Q| S  P  F  E  G  F  P  Q  S  V  L  Q  I  P |V  S  A  A  S  S  T  L  V  S  I  F  Q|                    180

A1   AGAGTTCCCCTGAGAGTACT CAAAGTCCTTTTGAGGGGTTT TCCCCAGTCTCCACTCCAGA TTCCTGTGAGGCCGCCGCCTTC TCCTCCACTTTATTGAGTAT   894
C1   |S  P  E  S  Q| S  P  F  E  G  F  P  Q  S  P  L  Q  I  P |V  S  R  S  F  S  S  T  L  V  S  I  F  Q|                    213

A1   TTTCCAGAGGTCCCCTGAGA GAAGTCAGAGAGAACTTCTGAG GGTTTTGCACAGTCTCCTCT CCAGATTCCTGTGAGCTCCT CCTCGTCCTCCACTTTACTG   994
C1   |F  Q  S  S  P  E  R  S  Q| |R  T  S  E  G  F  A  Q  S  P  L  Q  I  P| |V  S  S  S  S  S  T  L  L|                    246
```

FIG. 2(C)

C1 AGTCTCTTTCCAGAGTCCCC TGAGAGAACTCAGAGTACTT TTGAGGGTTTTCCCCAGTCT CCACTCCAGATTCCTGTGAG CCGGCTCCTTCCTTCTCCACTT 1094
A1 S _ L _ F _ Q _ S _ S _ P _ E _ R _ T _ Q | S _ T _ F _ E _ G _ F _ P _ Q _ S _ P _ L _ Q _ I _ P | V _ S _ _ R _ S _ F _ S _ S _ T _ L

C1 TATTGAGTAGTATTTTCCAGAGT TCCCCTGAGAGAGAACTCAGAG TACTTTTGAGGGGGTTTTGCCC AGTCTCCTCTCCAGATTCCT GTGAGCTCCTCCTTCTCCTCCTTC 1194
A1 _ _ L _ S _ I _ F _ Q _ S _ _ _ S _ P _ E _ R _ T _ Q _ _ S _ T _ F _ E _ G _ F _ A _ Q _ _ S _ P _ L _ Q _ I _ P | V _ S _ S _ S _ S _ S

C1 CACTTTATTGAGTAGTCTTTTCC AGAGTTCCCCTGAGAGAGAACT CAGAGTACTTTTGAGGGGGTTT TCCCCAGTCTCCTCTCCTTCTCCTCCAGA TTCCTATGACCTCCTCCTTC 1294
A1 _ T _ _ L _ L _ S _ L _ F _ Q _ _ _ S _ _ _ S _ P _ E _ R _ T _ Q | S _ T _ F _ E _ G _ F _ P _ Q _ S _ L _ L _ Q _ I _ P | M _ _ T _ S _ S _ F

C1 TCCCTCTACTTTATTGAGTAGTAT TTTCCAGAGTCTCCTGAGA GTGCTCAAAGTACTTTTGAG GGTTTTCCCCAGTCTCCTCT CCAGATTCCTGGGAGCCCCT 1394
A1 S _ S _ T _ _ L _ L _ S _ _ _ I _ F _ Q _ S _ S _ _ _ S _ A _ Q | S _ T _ F _ E _ G _ F _ P _ Q _ S _ P _ L _ Q _ I _ P | G _ S _ P _ S

C1 CCTTCTCCTCCCTCCACTTTACTG AGTCTTTTCCAGAGTCCCC TGAGAGAACTCACAGTACTT TCTCCTGAGAGTGCTCAAAG TGCTTTTGAGGGGGTTTTCCCC 1494
A1 _ F _ _ S _ S _ T _ _ L _ L _ _ _ S _ _ _ L _ F _ Q _ S _ _ _ S _ P _ E _ R _ T _ H | S _ T _ F _ E _ G _ F _ P _ Q _ S _ A _ E _ G _ F _ P _ Q

C1 CTCCCTCCTCCTCCTCTACTT TATTGAGTAGTATTTTACAGAGT CAGAGTACTTTTGAGGGGGTTT TCCCCAGTCTCCTCTCCTTCTCCTCCAGA 1594
A1 _ S _ S _ F _ S _ S _ _ _ _ L _ L _ S _ L _ F _ Q _ S _ _ _ S _ P _ E _ S _ A _ Q | S _ A _ F _ E _ G _ F _ P _ Q _ S _ P _ L _ Q _ I _ P

C1 GTGAGCTCCTCTTTCTCCCTA CACTTTATTGAGTAGTCTTTTCC AGAGTTCCCCTGAGAGAGAACT CAGAGTACTTTTGAGGGGGTT TCCCCAGTCTCCTCCTCCCTG 1694
A1 V _ S _ S _ F _ S _ Y _ _ _ L _ L _ S _ L _ F _ Q _ S _ _ _ S _ P _ E _ R _ T _ Q | S _ T _ F _ E _ G _ F _ P _ Q _ S _ P _ L _ Q

C1 TTCCCTGTGTGAGCTCCTCCTCC TCCTCCTCCTCCACTTTATTGAG TCTTTTCCAGAGTCCCCTG AGTGTACTCAAAGTACTTTT GAGGGTTTTCCCCAGTCTCC 1794
A1 _ P | V _ S _ S _ S _ S _ S _ _ _ S _ S _ T _ _ L _ L _ S _ _ _ L _ F _ Q _ S _ S _ P _ E _ _ _ C _ T _ Q | S _ T _ F _ E _ G _ F _ P _ Q _ S _ P 513

C1 1280
C1 1313
C1 1346
C1 1380
C1 1413
C1 1446
C1 1480

FIG. 2(D)

```
A1
C1  TCTCCCAGATTCCTCAGAGTC  CTCCTGAAGGGGGGAGAATACC  CATTCTCCTCCTCCAGATTGT  TCCAAGTCTTCCTGAGTGGG  AGGACTCCCTCCGTCTCCTCAC   1894
    L  Q  I  P  Q  S  P  P  E  G  E  N  T  H  S  P  L  Q  I  V  P  S  L  P  E  W  E  D  S  L  S  P  H          546

A1
C1  TACTTTCCTCAGAGCCCTCC  TCAGGGGGGAGGAGGACTCCCTAT  CTCCTCACTACTTTCCTCAG  AGCCCCTCCAGGGGGGAGGA  CTCCCTCCGTCTCCTCACTACT   1994
    Y  F  P  Q  S  P  P  Q  G  E  D  S  L  S  P  H  Y  F  P  Q  S  P  P  Q  G  E  D  S  L  S  P  H  Y  F       580
    P  Q  S  P  Q  G  E  D  S

A1
C1  TTCCCTCAGAGCCCTCAGGGG  GAGGACTCCCTGTCCGTCTCCTCA  CTACTTTCCTCAGAGCCCCTC  CTCAGGGGGGAGGAGGACTCCATG  TCTCCTCTCTACTTCCCTCA   2094
    P  Q  S  P  Q  G  E  D  S  L  S  P  H  Y  F  P  Q  S  P  P  Q  G  E  D  S  M  S  P  L  Y  F  P  Q        613

A1
C1  GAGTCCCTCTTCAGGGGGGGAGG  AATTCCAGTCTTCTCTCCAG  AGCCCCTGTGAGCATCTGCTC  CTCCTCCACTCCACATCCAGTC  TTCCCCAGAGTTCCCTGAG   2194
    S  P  L  Q  G  E  E  F  Q  S  S  L  Q  S  P  V  S  I  C  S  S  S  T  P  S  S  L  P  Q  S  F  P  E        646

A1
C1  AGTTCTCAGAGTCCTCCTGA  GGGGGCCTGTGTCCAGTCTTCCTC  CCATAGTCCTCCAGAGCCCT  CCTGAGGGGGGATGCACTCCCA  ATCTCCTCTCCAGAGTCCTG   2294
    S  S  Q  S  P  P  E  G  P  V  Q  S  P  L  H  S  P  Q  S  P  E  G  M  H  S  Q  S  P  L  Q  S  P  E        680

A1
C1  AGAGTGGCTCCTGAGGGGGGAG  GATTCCCTGTCCGTCTCCTCCA  AATTCCTCAGAGTCCCTCTG  AGGGAGAGGAGGACTCCCTGTCT  TCTCTCCATTTCCTCCAGAG   2394
    S  A  P  E  G  E  D  S  L  S  P  L  Q  I  P  S  P  L  E  G  E  D  S  L  S  L  H  F  P  Q  S             713

A1
C1  TCCCTCCTGAGTGGGGGGAGGACT  CCCTCTCCTCTCCACTCTT  CCTCAGTTCCTCCTCCAGGG  GGAGGACTTCCAGTCTTCTCTC  TCCAGAGTCCTGTGAGTATC   2494
    P  P  E  W  E  D  S  L  S  P  L  H  F  F  P  P  Q  G  E  D  F  Q  S  S  L  Q  S  P  V  S  I           746

A1
C1  TGCTCCTCCTCCTCCACTTCTTT  GGAGTCTTCCCCAGAGTTTCC  CTGAGAGTCCTCAGAGTCCTCAGA  GTCCTCCTGAGGGGGGGCCTGCC  CAGTCTCCTCCAGAGTCC   2594
    C  S  S  S  T  S  L  S  L  P  Q  S  F  P  E  S  P  Q  S  P  E  G  P  A  Q  S  P  L  Q  R  P  V  S       780

A1
C1  GCTCCTCTTCTTCCTCTACACT  TTAGCGAGTCTTCTCCAAAG  TTCCCATGAGAGTCCTGAGA  GTCCTCCTGAGGGGGGGCCTGCC  CAGTCTCCTCCAGAGTCC   2694
    S  F  F  S  Y  T  L  A  S  L  L  Q  S  H  E  S  P  Q  S  P  E  G  P  A  Q  S  P  L  Q  S  P          813
```

FIG. 2(E)

```
A1  .................................................................
C1  TGTGAGCTCCTTCCCCCTCCT  CCACTTCATCGAGAGTCTTTCC  CAGAGTCTTCCTGTGAGCTC  CTTCCCCCTCCTCCCACTTCAT  CGAGTCTTTCCAAGAGTCTCC .........  2794
    [V - S - - F - P - S - S - _ - T - S - S - L - S - _ - Q - S - S - P|[V - S - - F - P - S - S - _ - T - _ - S - S - L - _]

A1  .................................  AGTGAGGGTCCAGCAG  CCGTGAAGAGGAGGGGGCCAA  89
C1  CCTGAGAGTCCTCTCCCAGAG  TCCTGTGATCTCCTCTCCT  CCTCCACTTCATTGAGCCCA  TTCAGTGAAGAGTCCAGCAG  CCCAGTAGATGAATATACAA  468
    P - E - S - P - L - Q - S  P|[V - _ - I - S - F - S - S - _ - S - _ - T - _ - S - L - S - P - _ - _S]  S - E - G - S - S - S  R - E - E - E - G - P - S

A1  GCACCCTCTTGTATCCTG...  .................................  GAGTCCTTGTTCCTGGAGCA  GTAATCACTAAGAAGGTGGC  2894
C1  GTTCCTCAGACACCCTTGCTA  GAGAGTGATTCCTTGACAGA  CAGCGAGTCCTTGATAGAGA  GCGAGCCCCTTGTTCACTTAT  ACACTGGATGAAAAGGTGGA  523
    T - S - C - I - L - - - .  .................................  E - S - L - F - R - A - V - I - T - K - K - V - A  108

A1  .........  .................................  .........  ...GAGTCCTTGTTCCTGGAGCA  GTAATCACTAAGAAGGTGGC  880
C1  CGAGTTGGGCGGCGGTTTCTC  TCCTCAAATATCAAGTGAAG  CAGCCTATCACAAAGGCAGA  GATGCTGACGAATGTCATCA  GCGAGGTACACGGGCTACTTT  3094
    D - L - V - G - F - L - L  L - K - Y - Q - V - K  Q - P - I - T - K - A - E  M - L - T - N - V - I - S  R - Y - T - G - Y - F  946

A1  TGATTTGGTTGGTGTTTTCTGC  .................................  CCGGAGCCAGTCACAAAGGCAGA  AATGCTGGAGAGTGTCATCA  AAAATTACAAGCACTGTTTT  140
C1  GTTCCTCAGACACCCTTGCTA  GAGAGTGATTCCTTGACAGA  CAGCGAGTCCTTGATAGAGA  GCGAGCCCCTTGTTCACTTAT  ACACTGGATGAAAAGGTGGA  623
    E - L - A - R - F - L - L  L - K - Y - Q - V - K  Q - P - I - T - K - A - E  M - L - E - S - V - I - K  N - Y - K - H - C - F  140

A1  CCTGAGAGTCCTCTCGGCAAAGC  CTCTGAGTCGCTTGCAGCTGG  TCTTTGGCATTGACGTGAAG  GAAGCAGACCCCACCGGCCA  CTCCTATGTCCTTGTCACCT  174
C1  CGAGTTGGGCGGCGGTTTCTC  TCCTCAAATATCAAGTGAAG  CAGCCTATCACAAAGGCAGA  GATGCTGACGAATGTCATCA  GCGAGGTACACGGGCTACTTT  723
    P - E - I - F - G - K - A  S - E - S - L - Q - L - V  F - G - I - D - V - K  E - A - D - P - T - G - H  S - Y - V - L - V - T - C  174

A1  CCTGTGATCTTCAGGAAAGC  CCGTGAGTTCATAGAGATAC  TTTTTGGGCATTTCCCTGAGA  GAAGTGGACCCT---GATGA  CTCCTATGTCTTTGTAAACA  3191
C1  ...........................  .................................  .................................  .................................  .................................  979
    P - V - I - F - R - K - A  R - E - F - I - E - I - L  F - G - I - S - L - R  E - V - D - P - - D  S - V - F - V - N - T  979
```

(sequence alignment figure — nucleotide and amino acid sequences for A1 and C1)

FIG. 2(F)

```
      Q   D   L   V   Q   E   K   Y   L   E   Y   R   Q   V   P   D   S   D   P   A   R   Y   E   F   L   W   G   P   R   A   L   A   E   T
A1  CAAGAITTGGGTGCAGGAAAA   GTACCTGGAGTACCGGCAGG   TGCCGGACAGTGACCCGGCA   CGCTATGAGTTCCTGTGGGG   TCCAAGGGCCCCTCGCTGAAA  1023
C1  AAAGTTTGGGGTGCAGGAACA   TTACCTAGAGTACCGGGAGG   TGCCCAACTCTTCCCCTCCT   CGTTACGAATTCCTGTGGGG   TCCAAGAGCTCATTCAGAAG   3491
      K   V   W   V   Q   E   H   Y   L   E   Y   R   E   V   P   N   S   S   P   P   R   Y   E   F   L   W   G   P   R   A   H   S   E   V  1079
                                                                                                                                   274

      S   Y   V   K   V   L       E   Y   V   I   K   V   S       A   R   V   R   F   F   F   P   S   L   R   E   A   L   R   E   E   E
A1  CCAGCTATGTGAAAGTCCTT   GAGTATGTGATCAAGGTCAG   TGCAAGAGTTCGGCTTTTTCT   TCCCATCCCTGCGTGAAGCA   GCTTTGAGAGAGGAGGAAGA   1123
C1  TCATTAAGAGGAAAGTAGTA   GAGTTTTTGGCCATGCTAAA   TTCCATCCTCTTACAAGGAT   GCTTTGAAAGATGTGGAAGA   3591
      I   K   R   V   V       E   F   L   A   M   L   K       N   T   V   P   I   T   F   P   S   S   Y   K   D   A   L   K   D   V   E   E  1112
                                                                                                                                   307

      G   V   OPA
A1  GGGGAGTCTGAGCATGAGTTG   CAGCCAAGGCCGGTGCACCT   TCCAGGGCCGGTGCCCAGCAG   CTTCCCCTGCCTCGTGTGAC   309
C1  GAGAGCCAGGCCCATAATTG   ACACCACAGATGATTCGACT   CAGTGTCATGTCCCCCCAGCT   TCTCTTCTGAGTGAAGTCTA   3691
      R   Q   A   I   I   D       T   D   D   S   T       A   T   E   S   A   S   S       S   V   M   S   P   S   F       S   S   E   OPA  1142
                                                                                                                                   1223

A1  ---ATGAGGGCCCATTCTTCA   CTCTGAAGAGAGCGGTCAGT   GTTCTCAGTAGTAG-                                           GTTTC              1279
C1  GGGGCAGAATTCTTCCCTCTGA   GTTTGAAGGGGGGGGCAGTCGAG   TTTCTACGTGGTGGAGGGGCC   TGGTTGAGGCTGGAGAGAAC   ACAGTGCTATTTGCAATTCT   3791
```

[figure — nucleotide/amino-acid sequence alignment of clones A1 and C1]

FIG. 3

FIG.4

Gene MAGE-C1

Gene MAGE-C2

Gene MAGE-A1

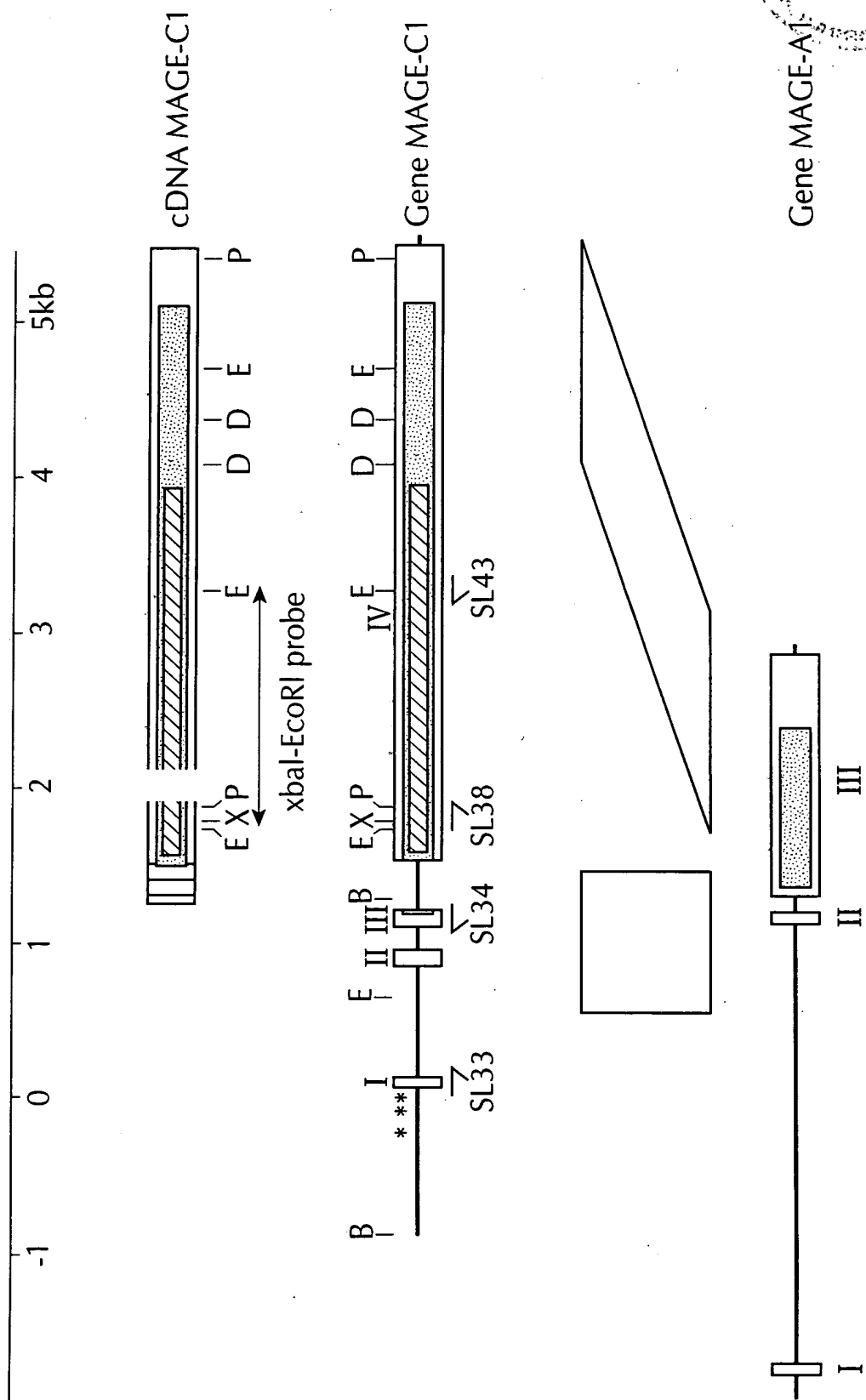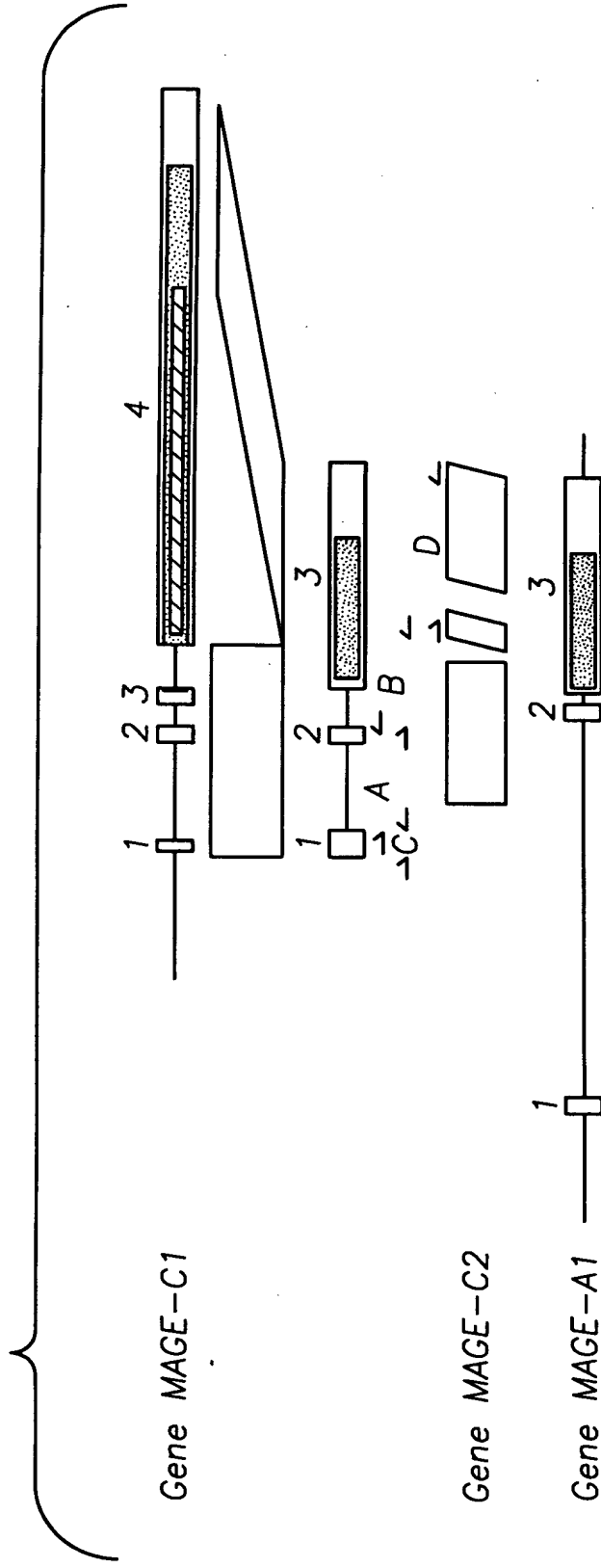**Figure 5**   **Nucleotide sequence of gene MAGE-C3**

(SEQ ID NO: 21)

The sequence corresponds to nucleotides 3761 to 4801 of PAC clone 232G24 (GenBank accession number AL022152). It has been renumbered in this figure, taking as nucleotide number 1 the first nucleotide of the start codon.

Primers SL164 and SL165 that were used in the RT-PCR assay are indicated.

```
ATGCCTCTCTTTCCAAACCTTCCACGCCTCAGCTTTGAGGAAGACTTCCA        50

GAACCCGAGTGTGACAGAGGACTTGGTAGATGCACAGGATTCCATAGATG       100

AGGAGGAGGAGGATGCCTCCTCCACTTCCTCTTCCTCTTTCCACTTTTTA       150

TTCCCCTCCTCCTCTTCCTTGTCCTCATCCTCACCCTTGTCCTCACCCTT       200
                        SL164
ACCCTCTACTCTCATTCTGGGTGTTCCAGAAGATGAGGATATGCCTGCTG       250

CTGGGATGCCACCTCTTCCCCAGAGTCCTCCTGAGATTCCTCCCCAGGGT       300

CCTCCCAAGATCTCTCCCCAGGGTCCTCCGCAGAGTCCTCCCCAGAGTCC       350

TCTAGACTCCTGCTCATCCCCTCTTTTGTGGACCCGATTGGATGAGGAGT       400

CCAGCAGTGAAGAGGAGGATACAGCTACTTGGCATGCCTTGCCAGAAAGT       450

GAATCCTTGCCCAGGTATGCCCTGGATGAAAAGGTGGCTGAGTTGGTGCA       500

GTTTCTTCTCCTCAAATATCAAACAAAAGAGCCTGTCACAAAGGCAGAGA       550

TGCTGACGACTGTCATCAAGAAGTATAAGGACTATTTTCCCATGATCTTC       600

GGGAAAGCCCATGAGTTCATAGAGCTAATTTTTGGCATTGCCCTGACTGA       650

TATGGACCCCGACAACCACTCCTATTTCTTTGAAGACACATTAGACCTCA       700

CCTATGAGGGAAGCCTGATTGATGACCAGGGCATGCCCAAGAACTGTCTC       750
                  SL165
CTGATTCTTATTCTCAGTATGATCTTCATAAAGGGCAGCTGTGTCCCCGA       800

GGAGGTCATCTGGGAAGTGTTGAGTGCAATAGGGGTGTGTGCTGGGAGGG       850

AGCACTTTATATATGGGGATCCCAGAAAGCTGCTCACTATACATTGGGTG       900

CAGAGAAAGTACCTGGAGTACCGGGAGGTGCCCAACAGTGCTCCTCCACG       950

TTATGAATTTTTGTGGGGTCCAAGAGCCCATTCAGAGGCCAGCAAGAGAA      1000

GTCTTAGAGTTTTTATCCAAGCTATCCAGTATCATCCCTAG             1041
```

Figure 6   Amino-acid sequence of the putative MAGE-C3 protein   (SEQ ID NO: 22)

```
MPLFPNLPRLSFEEDFQNPSVTEDLVDAQDSIDEEEEDASSTSSSSFHFL      50
FPSSSSLSSSSPLSSPLPSTLILGVPEDEDMPAAGMPPLPQSPPEIPPQG     100
PPKISPQGPPQSPPQSPLDSCSSPLLWTRLDEESSSEEEDTATWHALPES     150
ESLPRYALDEKVAELVQFLLLKYQTKEPVTKAEMLTTVIKKYKDYFPMIF     200
GKAHEFIELIFGIALTDMDPDNHSYFFEDTLDLTYEGSLIDDQGMPKNCL     250
LILILSMIFIKGSCVPEEVIWEVLSAIGVCAGREHFIYGDPRKLLTIHWV     300
QRKYLEYREVPNSAPPRYEFLWGPRAHSEASKRSLRVFIQAIQYHP        346
```

**Figure 7   Nucleotide sequence of gene MAGE-B5
(SEQ ID NO: 23)**

This sequence is the antiparallel sequence corresponding to nucleotides 123358 to 124185 of BAC clone 46604 (GenBank accession number AC005297). The sequence has been renumbered in this figure, taking the first nucleotide of the start codon as nucleotide number 1.

Primers SL189 and SL190 that were used in the RT-PCR assay are indicated.

```
ATGACTTCTGCAGGTGTTTTTAATGCAGGATCTGACGAAAGGGCTAACAG      50

TAGAGATGAGGAGTACCCATGTTCCTCAGAGGTCTCACCCTCCACTGAGA     100

GTTCATGCAGCAATTTCATAAATATTAAGGTGGGTTTGTTGGAGCAGTTC     150

CTGCTCTACAAGTTCAAAATGAAACAGCGTATTTTGAAGGAAGATATGCT     200

GAAGATTGTCAACCCAAGATACCAAAACCAGTTTGCTGAGATTCACAGAA     250

GAGCTTCTGAGCACATTGAGGTTGTCTTTGCAGTTGACTTGAAGGAAGTC     300

AACCCAACTTGTCACTTATATGACCTTGTCAGCAAGCTGAAACTCCCCAA     350

CAATGGGAGGATTCATGTTGGCAAAGTGTTACCCAAGACTGGTCTCCTCA     400
                        SL189
TGACTTTCCTGGTTGTGATCTTCCTGAAAGGCAACTGTGCCAACAAGGAA     450

GATACCTGGAAATTTCTGGATATGATGCAAATATATGATGGGAAGAAGTA     500

CTACATCTATGGAGAGCCCAGGAAGCTCATCACTCAGGATTTCGTGAGGC     550

TAACGTACCTGGAGTACCACCAGGTGCCCTGCAGTTATCCTGCACACTAT     600

CAATTCCTTTGGGGTCCAAGAGCCTATACTGAAACCAGCAAGATGAAAGT     650

CCTGGAATATTTGGCCAAGGTCAATGATATTGCTCCAGGTGCCTTCTCAT     700
                                SL190
CACAATATGAAGAGGCTTTGCAAGATGAGGAAGAGAGCCCAAGCCAGAGA     750

TGCAGCCGAAACTGGCACTACTGCAGTGGCCAAGACTGTCTCAGGGCGAA     800

GTTCAGCAGCTTCTCTCAACCCTATTGA     828
```

Figure 8   Amino-acid sequence of the putative MAGE-B5 protein (SEQ ID NO: 24)

```
MTSAGVFNAGSDERANSRDEEYPCSSEVSPSTESSCSNFINIKVGLLEQF      50
LLYKFKMKQRILKEDMLKIVNPRYQNQFAEIHRRASEHIEVVFAVDLKEV     100
NPTCHLYDLVSKLKLPNNGRIHVGKVLPKTGLLMTFLVVIFLKGNCANKE     150
DTWKFLDMMQIYDGKKYYIYGEPRKLITQDFVRLTYLEYHQVPCSYPAHY     200
QFLWGPRAYTETSKMKVLEYLAKVNDIAPGAFSSQYEEALQDEEESPSQR     250
CSRNWHYCSGQDCLRAKFSSFSQPY                              275
```

Figure 9   Nucleotide sequence of gene MAGE-B6
(SEQ ID NO: 25)

This sequence is the antiparallel sequence corresponding to nucleotides 146417 to 147640 of BAC clone 46604 (GenBank accession number AC005297). The sequence has been renumbered in this figure, taking the first nucleotide of the start codon as nucleotide number 1.

Primers SL191 and SL192 that were used in the RT-PCR assay are indicated.

```
ATGCCTCGGGGTCACAAGAGTAAGCTCCGTACCTGTGAGAAACGCCAAGA      50
GACCAATGGTCAGCCACAGGGTCTCACGGGTCCCCAGGCCACTGCAGAGA     100
AGCAGGAAGAGTCCCACTCTTCCTCATCCTCTTCTCGCGCTTGTCTGGGT     150
                 SL191
GATTGTCGTAGGTCTTCTGATGCCTCCATTCCTCAGGAGTCTCAGGGAGT     200
GTCACCCACTGGGTCTCCTGATGCAGTTGTTTCATATTCAAAATCCGATG     250
TGGCTGCCAACGGCCAAGATGAGAAAAGTCCAAGCACCTCCCGTGATGCC     300
TCCGTTCCTCAGGAGTCTCAGGGAGCTTCACCCACTGGCTCTCCTGATGC     350
AGGTGTTTCAGGCTCAAAATATGATGTGGCTGCCAACGGCCAAGATGAGA     400
AAAGTCCAAGCACTTCCCATGATGTCTCCGTTCCTCAGGAGTCTCAGGGA     450
GCTTCACCCACTGGCTCGCCTGATGCAGGTGTTTCAGGCTCAAAATATGA     500
TGTGGCTGCCGAGGGTGAAGATGAGGAAAGTGTAAGCGCCTCACAGAAAG     550
                 SL192
CCATCATTTTTAAGCGCTTAAGCAAAGATGCTGTAAAGAAGAAGGCGTGC     600
ACGTTGGCGCAATTCCTGCAGAAGAAGTTTGAGAAGAAAGAGTCCATTTT     650
GAAGGCAGACATGCTGAAGTGTGTCCGCAGAGAGTACAAGCCCTACTTCC     700
CTCAGATCCTCAACAGAACCTCCCAACATTTGGTGGTGGCCTTTGGCGTT     750
GAATTGAAAGAAATGGATTCCAGCGGCGAGTCCTACACCCTTGTCAGCAA     800
GCTAGGCCTCCCCAGTGAAGGAATTCTGAGTGGTGATAATGCGCTGCCGA     850
AGTCGGGTCTCCTGATGTCGCTCCTGGTTGTGATCTTCATGAACGGCAAC     900
TGTGCCACTGAAGAGGAGGTCTGGGAGTTCCTGGGTCTGTTGGGGATATA     950
TGATGGGATCCTGCATTCAATCTATGGGGATGCTCGGAAGATCATTACTG    1000
AAGATTTGGTGCAAGATAAGTACGTGGTTTACCGGCAGGTGTGCAACAGT    1050
GATCCTCCATGCTATGAGTTCCTGTGGGGTCCACGAGCCTATGCTGAAAC    1100
CACCAAGATGAGAGTCCTGCGTGTTTTGGCCGACAGCAGTAACACCAGTC    1150
CCGGTTTATACCCACATCTGTATGAAGACGCTTTGATAGATGAGGTAGAG    1200
AGAGCATTGAGACTGAGAGCTTAA                              1224
```

Figure 10 **Amino-acid sequence of the putative MAGE-B6 protein (SEQ ID NO: 26)**

```
MPRGHKSKLRTCEKRQETNGQPQGLTGPQATAEKQEESHSSSSSSRACLG        50
DCRRSSDASIPQESQGVSPTGSPDAVVSYSKSDVAANGQDEKSPSTSRDA       100
SVPQESQGASPTGSPDAGVSGSKYDVAANGQDEKSPSTSHDVSVPQESQG       150
ASPTGSPDAGVSGSKYDVAAEGEDEESVSASQKAIIFKRLSKDAVKKKAC       200
TLAQFLQKKFEKKESILKADMLKCVRREYKPYFPQILNRTSQHLVVAFGV       250
ELKEMDSSGESYTLVSKLGLPSEGILSGDNALPKSGLLMSLLVVIFMNGN       300
CATEEEVWEFLGLLGIYDGILHSIYGDARKIITEDLVQDKYVVYRQVCNS       350
DPPCYEFLWGPRAYAETTKMRVLRVLADSSNTSPGLYPHLYEDALIDEVE       400
RALRLRA                                                  407
```